Chapter 6
Modeling Human Perceptions in e-Commerce Applications: A Case Study on Business-to-Consumers Websites in the Textile and Fashion Sector

Adrian Castro-Lopez and Jose M. Alonso

6.1 Introduction

The fashion market is global and presents a complex structure, which operates in many different levels to reach all kinds of public, from those who love fashion to those who believe that buying clothes is a daily necessity. Fashion is a global industry with a market value of around $1.7 billion [17].

Besides, the textile and fashion sector via Internet is placed among those with higher importance revenue figures in the worldwide on-line market [22]. The main sectors of activity in Spain are detailed in [10]: Travel agents and tour operators (14.4%), Flights (11.9%) and Clothing (5.4%). Moreover, the remarkable growth rates in the clothing sector in recent years have led fashion companies to use the Business-to-Consumers (B2C) on-line channel as a mean for promoting and selling.

The success or failure of different B2C websites highly depends on the e-service quality perceived by consumers. The e-service quality can be defined as [16]:

> The extent to which a website provides effective and efficient results in regard to the information search process, to the purchase and delivery of products and services, and even to the client enjoyment and emotional experience.

In this regard, there are several models of e-service quality (e.g., ESQ and Customer Experience [20] or New PeSQ [19]). To sum up with, previous studies

A. Castro-Lopez
Business Administration Department, Faculty of Economic and Business, University of Oviedo, Oviedo, Spain
e-mail: castroadrian@uniovi.es

J. M. Alonso (✉)
Centro Singular de Investigación en Tecnoloxías da Información (CiTIUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain
e-mail: josemaria.alonso.moral@usc.es

A. Meier et al. (eds.), *Applying Fuzzy Logic for the Digital Economy and Society*, Fuzzy Management Methods, https://doi.org/10.1007/978-3-030-03368-2_6

have already identified different latent dimensions in e-service quality. In addition, they suggest that it is necessary to analyze hedonic and utilitarian dimensions of e-service quality. Utilitarian quality is defined as the value derived from completing objectives, from finding information, and/or from buying. Hedonic quality is defined as the value derived from enjoying the search for information and/or for purchasing. Moreover, many e-service quality models rely on inquiries to consumers about their perception on both hedonic and utilitarian dimensions. However, consumer perceptions on qualitative issues are likely to suffer from high levels of uncertainty and vagueness. Therefore, it is required to find a suitable methodology to deal with the uncertainty inherent to consumer perceptions in e-service quality assessment and modeling.

Fuzzy Logic provides a framework to the Computational Theory of Perceptions (CTP) [24] which is acknowledged for its well-known ability for approximate reasoning and linguistic concept modeling; mainly due to its semantic expressivity close to natural language. Fuzzy sets and systems are able to mathematically formalize, in an approximate but even precise way, uncertainty and vague concepts (like hedonic and utilitarian dimensions of e-service quality). In addition, interpretability of fuzzy sets and systems [2], due to its human-centric character, plays a key role in system modeling and it becomes essential in applications with high human interaction like sensory evaluation [25]. Moreover, a recent survey on the eXplainable Artificial Intelligence (XAI) research field [3] has shown the relevance of interpretable fuzzy systems in the quest for XAI systems. Notice that, the recent success of many AI applications into real-world usage has triggered some critical voices regarding ethical and legal issues. Moreover, the new European General Data Protection Regulation (GDPR[1]), approved by the European Parliament and to take effect in May 2018, refers to the "right to explanation" to European citizens. This new regulation makes even more appealing the design of XAI systems in general, and the modeling of interpretable fuzzy systems in particular, as a way to pave the way towards XAI.

The purpose of this study is to expand and further explore the knowledge on e-service quality. We combine marketing methods (qualitative and quantitative methods) and CTP (Fuzzy Logic) for the assessment and modeling of e-service quality. As a result, we get a more dynamic evaluation, enhancing adaptability to changing needs of consumer perceptions. Accordingly, business managers can redirect the investment strategies and focus on what is actually valued by consumers. Thus, in this paper we contribute to the field of analysis on e-service quality as follows:

- Data acquisition is addressed in terms of collecting consumer perceptions (regarding hedonic and utilitarian dimensions) through fuzzy rating scale-based questionnaires [18].

[1] http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32016R0679.

- Collected data are processed under the fuzzy logic formalism provided by CTP. Thus, we deal efficiently with uncertainty and vagueness all along the processing chain, including aggregation and fusion in the search of consensus agreement among groups of consumers. Notice that we consider the use of fuzzy multi-criteria decision-making tools [21].
- The relation between the main dimensions of e-service quality is modeled by means of a set of linguistic variables and fuzzy IF-THEN rules. The result is an interpretable fuzzy system which combines both expert knowledge and knowledge automatically extracted from data [4].
- The proposal is validated with a study of the main Business-to-Consumers websites in the Spanish textile and fashion sector.

The rest of the paper is structured as follows. Section 6.2 introduces the materials and methods applied to carry out this study. Then, Sect. 6.3 presents and discusses the main reported results. Finally, main conclusions and future perspectives are sketched in Sect. 6.4.

6.2 Materials and Methods

6.2.1 Survey Methods

Survey methods have been worldwide applied to collect opinions from consumers. Surveys supported by Likert scales [14] are likely to be the most usual ones, mainly because of their simplicity. Respondents (also called assessors in the field of sensory sciences) are usually asked to choose an answer among a small set of options (commonly expressed by ordered linguistic terms). This fact implies a lack of flexibility that is argued as the main disadvantage of this kind of surveys [6]. Moreover, the goodness of drawn conclusions strongly depends on how carefully surveys were designed in order to avoid bias and minimize ambiguity, imprecision and uncertainty in the given questionnaires. Of course, understanding properly the meaning of the involved linguistic terms depends on the context and background of each respondent. In addition, human perceptions and opinions are always subjective and it is not feasible to check how truthful respondents are. From a psychometric point of view, fuzzy rating scales make easier the assessment of the diversity, subjectivity, imprecision and uncertainty which are inherent to human perceptions [11]. Surveys supported by fuzzy rating scale-based questionnaires are especially helpful in practical applications. For example, Gil et al. applied them to teaching evaluation [12]. Moreover, Quirós et al. proved their utility in relation with the customized packaging design of gin bottles [18].

6.2.2 Quantitative and Qualitative Analysis of Human Perceptions

In a previous study [18], we proposed a new methodology for descriptive and comparative analysis of human perceptions expressed through fuzzy questionnaires. The treatment of collected data requires the adaptation of statistical techniques to the fuzzy case. It is worthy to note that SMIRE[2] researchers have actively developed statistical tools around the concept of fuzzy rating scale [11, 12]. Moreover, they have provided the research community with the free software *R* package called SAFD.[3]

Both the design of a specific fuzzy questionnaire and the analysis of collected data are made with the Quale software [5]. Quale implements the methodology described in [18] and calls to SAFD for dealing with fuzzy statistics. Moreover, it produces as result a survey report made up of a set of graphs and texts in a user-friendly style which can be customized in accordance with the reader background and preferences. Firstly, sensory data acquired through fuzzy rating scale-based questionnaires are formalized under fuzzy logic formalism. The three values that characterize each given evaluation are translated into a triangular fuzzy set $A = (a; b; c; h)$, where b represents the modal point (upper value of the fuzzy triangle), a and c determine the support (lower confidence interval), and h is the height of the triangle (by default it takes value 1). Let X be a non-empty set. Being $FS(X)$ the set of all fuzzy sets in X, $A_i = (a_i, b_i, c_i) \in FS(X)$ corresponds to the evaluation provided by assessor i. Once a set of evaluations have been collected regarding a specific sample, then they are aggregated by the sample Aumann-type mean:

$$\frac{1}{n}\sum_{i=1}^{n} A_i = \left(\frac{1}{n}\sum_{i=1}^{n} a_i, \frac{1}{n}\sum_{i=1}^{n} b_i, \frac{1}{n}\sum_{i=1}^{n} c_i\right) \tag{6.1}$$

We group those points in the scale with the greatest aggregated values until a fixed threshold of the total is reached. Then, we build the intervals that best shape the set of points. In case two or more intervals are close enough, they are fused into a single interval. Later, we compute the center of gravity (COG) of the most representative interval (that one covering most evaluations). Given a triangular fuzzy set $A \in FS(X)$, COG is calculated as follows:

$$COG(A) = \min\{y \in [a, c] | \int_a^y \mu_A(x)\mathrm{d}x \geq 0.5\} \tag{6.2}$$

[2]SMIRE stands for Statistical Methods with Imprecise Random Elements. This is the name of the Statistics and Fuzzy Logic research group in the University of Oviedo (Spain).

[3]SAFD stands for Statistical Analysis of Fuzzy Data. This *R* package is available at https://cran.r-project.org/web/packages/SAFD/index.html [Accessed on May 2018].

where $\mu_A(x)$ measures the membership degree of x to A. $COG(A)$ represents the aggregated score associated to the sample and attribute under study. In addition, the number of evaluations characterized by a fuzzy set A is given by:

$$p_A = \sum_{i=1}^{m} S(A_i, A) \tag{6.3}$$

where $S(A_i, A)$ measures the degree up to which A_i is a subset of A:

$$S(A_i, A) = 1 - \frac{\sum_{x \in X} \max(0, \mu_{A_i}(x) - \mu_A(x))}{\sum_{x \in X} \mu_{A_i}(x)} \tag{6.4}$$

The samples under study are ranked with respect to their related scores. Those samples without faithful scores are set "in quarantine" and separated from the rest. We consider three situations which denote a lack of consensus:

- The main interval is too narrow. Thus, it does not characterize a big enough number of assessors.
- The main interval is too wide.
- There exists a second interval which becomes comparable to the main interval in terms of associated evaluations.

The interested reader is kindly referred to [18] for a deeper explanation of the Quale methodology that we have only sketched above for the sake of brevity.

6.2.3 Multi-Criteria Decision-Making Tools

There are different tools for the evaluation of a group of alternatives as a function of a finite number of criteria given by a decision maker or a group of them. Some of the basic methods are [13]: Weighted Sum Model (WSM), Weighted Product Model (WPM), Compromise Programming (CP), Analytical Hierarchy Process (AHP), Elimination and Choice Expressing Reality method (ELECTRE), Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), Preference Ranking Organization Method for Enrichment Evaluation (PROMETHEE) and Serbian Multi-criteria Optimization and Compromise Solution (VIKOR).

We will focus on the TOPSIS method and its fuzzy extension F-TOPSIS [23]. This is a suitable tool to handle properly the uncertainty that is intrinsic to the opinions in a decision-making process. F-TOPSIS has been successfully used in several applications (e.g., supply chain management [9] or shopping website evaluation [21]). The F-TOPSIS method is summarized in the next three steps:

- **Step 1.** Determination of the fuzzy decision matrix: Defining the n fuzzy evaluation criteria $(C_1,\ldots,C_j,\ldots,C_n)$ for all m alternatives $(A_1,\ldots,A_i,\ldots,A_m)$; and building the $m \times n$ matrix:

$$\left[\widetilde{D_x}\right] = \begin{pmatrix} \widetilde{x}_{11} & \cdots & \widetilde{x}_{1j} & \cdots & \widetilde{x}_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \widetilde{x}_{i1} & \cdots & \widetilde{x}_{ij} & \cdots & \widetilde{x}_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \widetilde{x}_{m1} & \cdots & \widetilde{x}_{mj} & \cdots & \widetilde{x}_{mn} \end{pmatrix} \tag{6.5}$$

  where $\widetilde{x}_{ij} = \left(\widetilde{x}^1_{ij}, \widetilde{x}^2_{ij}, \widetilde{x}^3_{ij}\right)$ is a triangular fuzzy number which corresponds to alternative A_i and criterion C_j, with $i \in [1, m]$ and $j \in [1, n]$.
- **Step 2.** Construction of the normalized and weighted decision matrix:

$$\widetilde{x}^*_{ij} = \left(\widetilde{x}^{1*}_{ij}, \widetilde{x}^{2*}_{ij}, \widetilde{x}^{3*}_{ij}\right); \widetilde{x}^{k*}_{ij} = \frac{\widetilde{x}^k_{ij}}{\max_j \left(\widetilde{x}^3_{ij}\right)}; \forall k \in [1, 3] \tag{6.6}$$

$$\begin{gathered} \widetilde{v}_{ij} = \widetilde{W}_j \times \widetilde{x}^*_{ij}; \widetilde{v}^k_{ij} = w^k_j \times \widetilde{x}^{k*}_{ij}; \widetilde{W}_j = \left(w^1_j, w^2_j, w^3_j\right) \\ \forall k \in [1, 3]; \forall i \in [1, m]; \forall j \in [1, n] \end{gathered} \tag{6.7}$$

- **Step 3.** Closeness coefficients for each alternative and ranking.
  – Determination of the Ideal Positive Fuzzy Solution ($FPIS^+$) and the Ideal Negative Fuzzy Solution ($FNIS^-$):

$$FPIS^+ = \left\{\widetilde{v}^+_1, \ldots, \widetilde{v}^+_j, \ldots, \widetilde{v}^+_n\right\}; \widetilde{v}^+_j = (1, 1, 1); \forall j \in [1, n] \tag{6.8}$$

$$FNIS^- = \left\{\widetilde{v}^-_1, \ldots, \widetilde{v}^-_j, \ldots, \widetilde{v}^-_n\right\}; \widetilde{v}^-_j = (0, 0, 0); \forall j \in [1, n] \tag{6.9}$$

  – Computing the distance between weighted criteria and the closeness coefficient for each alternative:

$$d^+_i = \sum_{i=1}^{n} d(\widetilde{v}_{ij}, \widetilde{v}^+_j); \forall i \in [1, m]; \forall j \in [1, n] \tag{6.10}$$

$$d^-_i = \sum_{i=1}^{n} d(\widetilde{v}_{ij}, \widetilde{v}^-_j); \forall i \in [1, m]; \forall j \in [1, n] \tag{6.11}$$

$$CC_i = \frac{d^-_i}{d^-_i + d^+_i}; \forall i \in [1, m] \tag{6.12}$$

The final ranking of alternatives is established in accordance with (6.12). The interested reader is kindly referred to [23] for a deeper explanation about F-TOPSIS.

6.2.4 Interpretable Fuzzy Modeling

Fuzzy techniques are ready to deal properly with imprecision and uncertainty in the identification and modeling of systems [7]. Namely, the Highly Interpretable Linguistic Knowledge (HILK) methodology [4] is aimed at designing fuzzy models by combining expert knowledge (derived by human knowledge-elicitation tasks such as interviews, surveys, and so on) and knowledge automatically extracted from data (derived by data-mining tasks). This methodology is implemented in the free software GUAJE[4] [15] which makes intuitive the generation of understandable and accurate fuzzy models.

A fuzzy model is made up of two main components: the knowledge base (KB) and the inference engine. On the one hand, the KB comprises a set of linguistic variables and rules (which combine expert and induced knowledge). Notice that knowledge representation tasks are carried out off-line. On the other hand, the inference engine is in charge of exploiting the model on-line.

Regarding the construction of the KB, a panel of experts is asked to define relevant variables and rules. In addition, we can apply data mining tools provided by GUAJE because the key issue in HILK is the careful combination of expert and induced knowledge. The entire modeling process comprises three steps:

- **Fuzzy partition design.** The goal is to define the most influential variables, according to both expert knowledge and knowledge extracted from data. On the one hand, experts provide complete or partial information about the identified variables. On the other hand, several algorithms can be used to create fuzzy partitions form data. The result is the definition of a common universe for each variable according to both expert knowledge and data distribution. Notice that linguistic constraints (distinguishability, normalization, coverage, overlapping, etc.) have to be superimposed to the fuzzy partition definition in order to ensure interpretability [2]. Thus, we recommend the use strong fuzzy partitions which satisfy all previous interpretability constraints and are defined as follows:

$$\sum_{i=1}^{M} \mu_{A_i}(x) = 1, \ \forall x \in U \tag{6.13}$$

  where U=[U_l,U_u] is the universe of discourse, M is the number of linguistic terms, and $\mu_{A_i}(x)$ is the membership degree of x to the A_i fuzzy set.

[4] http://sourceforge.net/projects/guajefuzzy/.

- **Rule base definition and integration.** Experts are invited to describe the system behavior through linguistic rules (*Expert Rules*). In addition, rules are induced from data (*Induced Rules*). Both types of rules use the same linguistic terms defined by the same fuzzy sets. Rule format is as follows:

$$\textbf{If} \quad \underbrace{\underbrace{X_a \text{ is } A_a^i}_{Partial\ Premise\ P_a} \quad \textbf{AND} \ldots \textbf{AND} \quad \underbrace{X_z \text{ is } A_z^j}_{Partial\ Premise\ P_z}}_{Premise} \quad \textbf{Then} \quad \underbrace{Y \text{ is } C^n}_{Conclusion}$$

- **KB improvement.** The goal of this step is to enhance the KB interpretability-accuracy trade-off. First, the KB quality is assessed according to both accuracy and interpretability. Second, a simplification procedure is run to increase interpretability without penalizing either consistency or accuracy. Third, an optimization process is applied to get better accuracy while keeping interpretability.

The interested reader is referred to [4] for more details about the HILK methodology. In addition, a thorough review on fuzzy system software is given in [1].

6.3 Results

This section goes in depth with the results coming out from applying the materials and methods previously introduced to a use case regarding B2C websites in the textile and fashion sector. For the sake of readability, the section is split into two additional ones. We start with presenting and discussing results related to e-service quality analysis. Then, we focus on results related to e-service quality modeling.

6.3.1 e-Service Quality Analysis

In a preliminary study [8], we carried out a classical Likert-based survey on the Spanish textile and fashion sector. We collected opinions of a sample of 405 habitual consumers from sales platforms. The survey was disseminated through social networks (Facebook, LinkedIn and Twitter), by email and through personal interviews. The sampling error was ±2.42% with a trust level of 95% (p = q = 0.5). The sample distribution was done by levels of age (21% between 18 and 24 years, 49% between 25 and 34 years, 19% between 35 and 44 years, 11% over 45 years) and gender (60% women, 40% men).

The questionnaire was made for two groups of consumers: (1) those consumers who only search for information (40%), and those ones who search for information

Table 6.1 Ranking of B2C websites provided by F-TOPSIS

	eBay		Zara		Privalia		Buy Vip		Vente Privee		Asos		El Corte Inglés	
	d^+	d^-	d^+	d^-	d^+	d^-	d^+	d^-	d^+	d^-	d^+	d^-	d^+	d^-
C_1	0.42	0.61	0.35	0.69	0.36	0.68	0.37	0.67	0.35	0.69	0.35	0.69	0.44	0.59
C_2	0.38	0.65	0.32	0.72	0.33	0.71	0.34	0.70	0.31	0.74	0.31	0.74	0.36	0.68
C_3	0.43	0.60	0.31	0.73	0.34	0.70	0.35	0.69	0.35	0.69	0.35	0.68	0.32	0.72
C_4	0.38	0.66	0.40	0.63	0.36	0.68	0.37	0.66	0.35	0.69	0.35	0.69	0.43	0.60
C_5	0.64	0.38	0.62	0.40	0.68	0.34	0.66	0.36	0.78	0.24	0.60	0.42	0.69	0.33
C_6	0.39	0.65	0.37	0.67	0.37	0.66	0.37	0.66	0.43	0.60	0.30	0.74	0.35	0.68
C_7	0.37	0.66	0.37	0.66	0.36	0.67	0.34	0.69	0.45	0.57	0.34	0.70	0.38	0.66
C_8	0.35	0.68	0.45	0.57	0.43	0.59	0.45	0.58	0.59	0.43	0.46	0.57	0.42	0.61
C_9	0.46	0.57	0.38	0.65	0.46	0.57	0.46	0.57	0.48	0.55	0.35	0.69	0.51	0.51
CC_i	**0.590**		**0.615**		**0.602**		**0.600**		**0.560**		**0.634**		**0.580**	
Ranking	**5**		**2**		**3**		**4**		**7**		1		**6**	

but also buy (60%). We asked about the B2C websites of the next seven retailers: *eBay*,[5] *Zara*,[6] *Privalia*,[7] *Buy Vip*,[8] *Vente Privee*,[9] *Asos*,[10] and *El Corte Inglés*.[11]

In the light of collected data, we first identified the following latent dimensions and factors (C_i) to consider when assessing e-Service Quality:

- *Utilitarian Quality*:
  - Website Quality: Design (C_1) and Contents (C_2).
  - Offered Service: Guarantee (C_3), Offer (C_4), and Customization (C_5).
  - Security: Payment management (C_6), Privacy (C_7), and Trust (C_8).
- *Hedonic Quality* (C_9).

Then, we applied F-TOPSIS (briefly introduced in Sect. 6.2.3) with the aim of ranking the seven B2C websites under study with respect to the nine C_j factors listed above. Table 6.1 summarizes the results of applying F-TOPSIS on the available data. This table is structured as follows. Columns are related to websites while rows are related to factors. For each website, we report positive (d^+) and negative (d^-) distance between the weighted criteria and ideal solutions. At the bottom, the last two rows show the closeness coefficients CC_i and the final ranking. The B2C website of *Asos* turns up with the highest score (0.634) for e-Service Quality. However, it is closely followed by *Zara* (0.615). Behind them, we find *Privalia*

[5] http://www.ebay.es/.
[6] http://www.zara.com.
[7] http://www.privalia.com/.
[8] http://es.buyvip.com/.
[9] http://www.vente-privee.com/.
[10] http://www.asos.com.
[11] https://www.elcorteingles.es/.

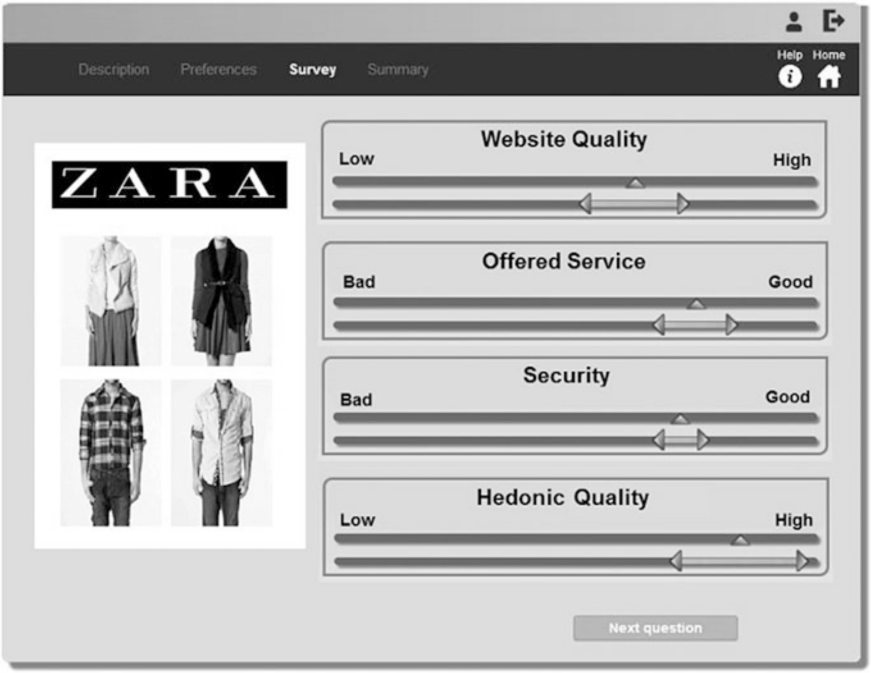

Fig. 6.1 Example of fuzzy rating scale-based questionnaire designed by Quale

(0.602), *Buy Vip* (0.600), *eBay* (0.590), *El Corte Inglés* (0.580), and *Vente Privee* (0.560).

Later, we designed a second survey with the aim of making a finer complementary study regarding the same seven retailers considered previously. This survey was supported by an on-line fuzzy rating scale-based questionnaire (see Fig. 6.1).

We collected data from 78 assessors. They were selected randomly, but respecting the same sample distribution, concerning those assessors who took part in the first study. For each website, assessors had to evaluate four attributes related to the main latent dimensions previously identified: (1) Website Quality, (2) Offered Service, (3) Security, and (4) Hedonic Quality. Notice that the first three attributes are related to the Utilitarian Quality. Each attribute was evaluated in a fuzzy rating scale like the ones depicted in Fig. 6.1. The narrower the triangle support, the more confident the answer is.

Both the design of the fuzzy questionnaire and the analysis of collected data were made as we briefly sketched in Sect. 6.2.2. As result, we obtained a report with, among others, the following contents:

- **Attribute correlation matrix.** We computed Pearson correlation between each pair of attributes under study. Figure 6.2 depicts the correlation matrix in the use case. As expected, the matrix is symmetrical. In addition, correlation is positive

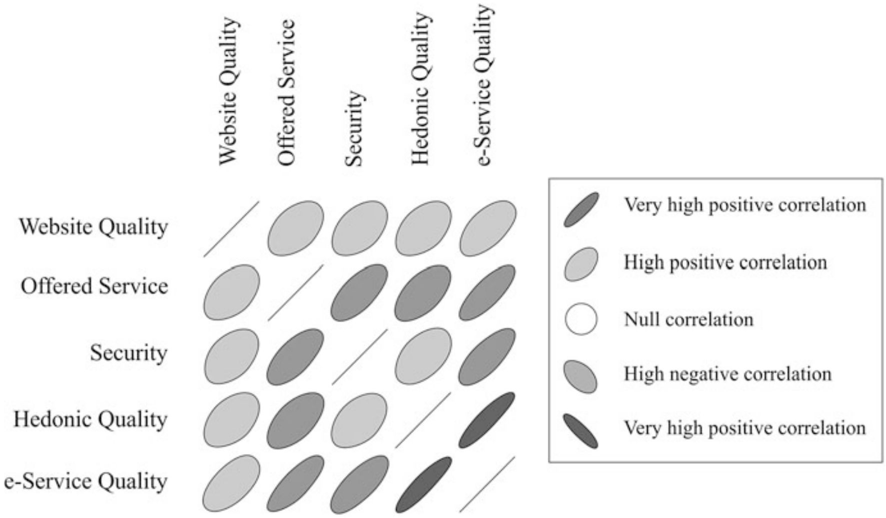

Fig. 6.2 Correlation matrix (Pearson)

in all cases. Moreover, it is easy to appreciate how e-Service Quality is mainly correlated with Hedonic Quality. With respect to the latent factors of Utilitarian Quality, we observe stronger correlation of e-Service Quality with the Offered Service and Security than with the Website Quality.

- **Spider plots.** These graphs summarize at once all collected assessments (in average score) regarding all attributes for a given sample. This fact makes intuitive the comparison among all websites under study (see Fig. 6.3).

  Each attribute is represented by a grey sector. The larger the area of the sector is, the higher the related score. All retailers get high score (above 74) for Website Quality, but the highest score (86) is achieved by *Vente Privee*. Nevertheless, *Vente Privee* gets the lowest score (66.75) regarding Offered Service. In addition, the best service is offered by *Zara* (79.25). From Security point of view, *eBay* is the most appreciated (77) while *Privalia* is the least appreciated (73.5). Notice that security of all websites is considered almost equal.

  As expected, the evaluation of Hedonic Quality exhibits a larger dispersion of answers and a smaller consensus. The highest score (94.5) is achieved by *Asos* while the lowest score (53.75) corresponds to *El Corte Inglés*.
- **Ranking of retailers regarding e-service quality.** Figure 6.4 shows a bar chart with all seven retailers under study. They are ordered in accordance with the average scores computed after processing the data collected in the second survey.

  On the left hand side of the picture, inside the rectangle, we can see bars which correspond to those websites for which assessors were in agreement. Among them, *Zara* (78.25) turns up as the one with the highest e-Service Quality, even though *Asos* (77) is not too far away. The lowest score corresponds to *El Corte Inglés* (although *Privalia* is close). It is worthy to note that we keep on the right

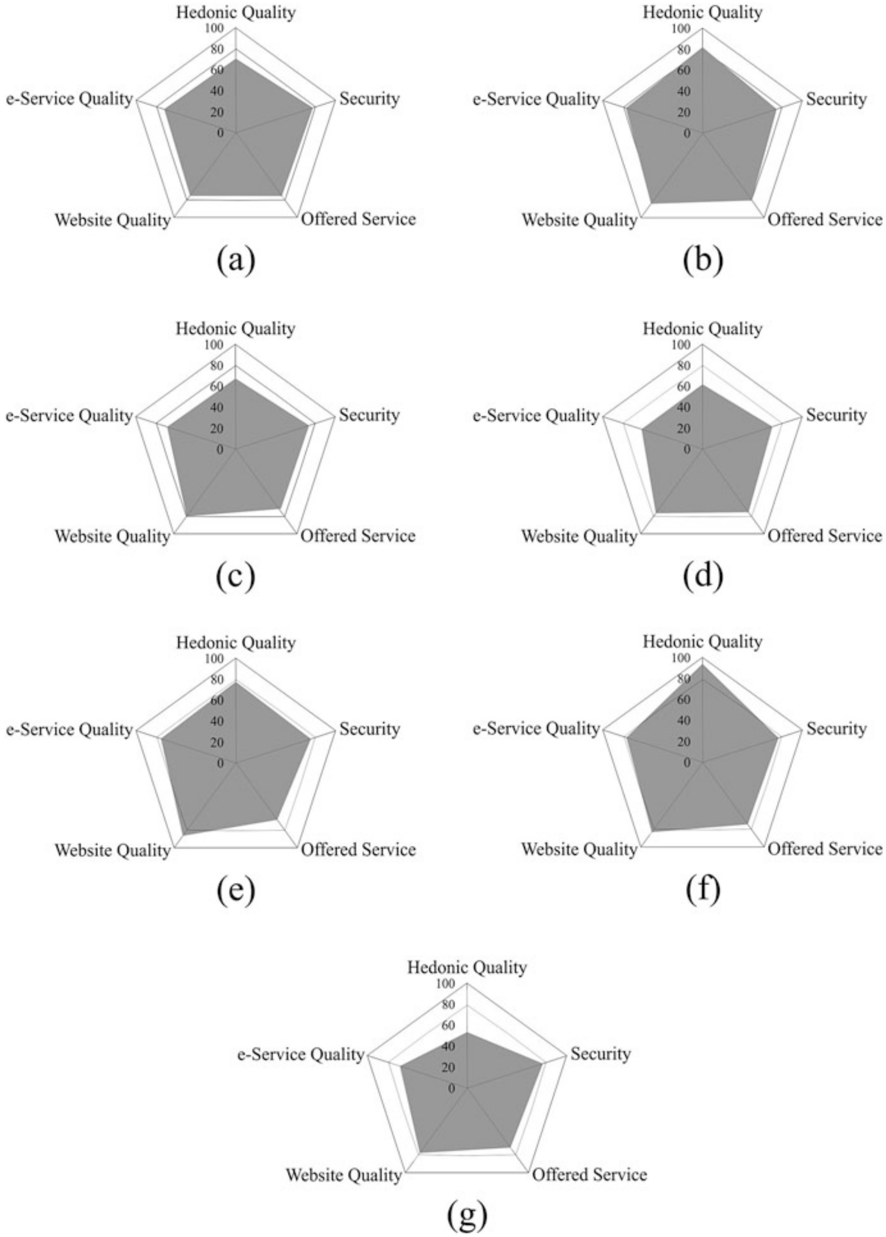

Fig. 6.3 Comparison among B2C websites by spider plots. (**a**) eBay. (**b**) Zara. (**c**) Privalia. (**d**) Buy Vip. (**e**) Vente Privee. (**f**) Asos. (**g**) El Corte Inglés

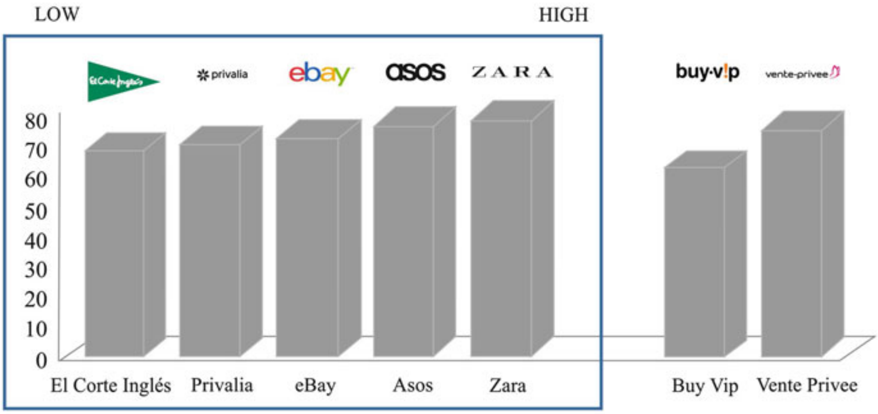

Fig. 6.4 Ranking of retailers provided by Quale with respect to e-service quality

side of the picture, out of the given ranking, the two retailers (*Buy Vip* and *Vente Privee*) for which assessors were not in agreement. So, their related score is not faithful and we must be careful in the comparison against the other retailers.

With the aim of giving a deeper insight with respect to the degree of consensus among assessors, it is needed to take a look at Fig. 6.5. It depicts the distribution of e-Service Quality assessments for all the seven retailers. In each picture, the horizontal axis shows the evaluation range [0,100] while the vertical axis yields the aggregated score normalized in [0,1]. On the one hand, the background curve characterizes all aggregated answers. On the other hand, the foreground bars identify the areas with the greatest answer accumulation. The height of each bar is proportional to the percentage of answers it covers (which is given on top of the bar).

As it can be seen in Fig. 6.5d, there are two disjoint bars to take care in the detailed analysis of e-Service Quality for *Buy Vip*. Moreover, the second bar (16%) is important enough in order not to be ignored. Quale remarks this fact through a warning symbol which is depicted as a triangle with an exclamation mark inside. Anyway, the main bar represents 55% of answers. Therefore, its center of gravity can be seen as a more representative score than the average score for the whole distribution. The situation is even worse in the case of *Vente Privee* (see Fig. 6.5e) where Quale yields a heavy warning (depicted as a double triangle with an exclamation mark inside) because the two bars are really close (39% versus 32%). This means we cannot trust on the aggregated score because we have two plausible values which are likely to yield to two different rankings. Thus, we recommend excluding *Vente Privee* from the final ranking which is as follows: *Zara* (78.25), *Asos* (77), *eBay* (72), *Privalia* (69.5), *El Corte Inglés* (68), and *Buy Vip* (63).

We would like to remark that this ranking is quite similar to the one provided by F-TOPSIS (see Table 6.1) but there are some subtle and valuable differences.

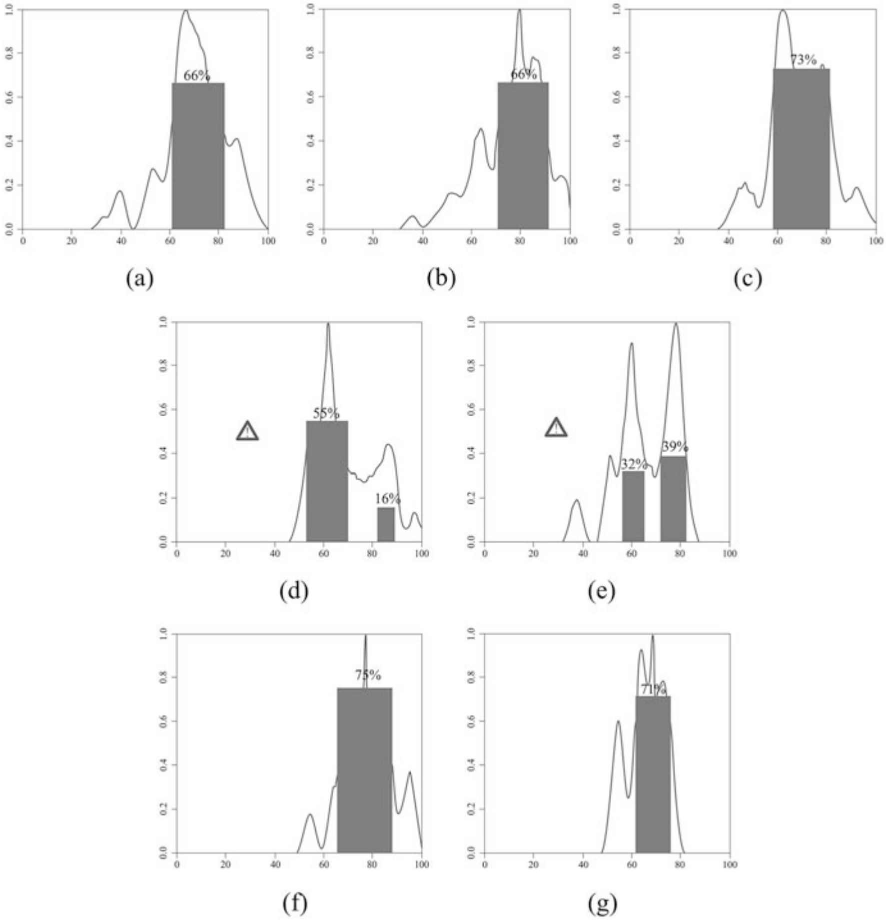

Fig. 6.5 Distribution of collected answers regarding e-service quality. (**a**) eBay. (**b**) Zara. (**c**) Privalia. (**d**) Buy Vip. (**e**) Vente Privee. (**f**) Asos. (**g**) El Corte Inglés

Firstly, both methods place *Zara* and *Asos* at the top of the ranking but with exchanged positions. Anyway, both retailers get so close scores that we can say there is not any difference between them. Secondly, far from the top, *eBay* and *Privalia* turn up also quite close in the middle of both rankings. In addition, *El Corte Inglés* is slightly behind and it goes to the last position in case of excluding the two retailers (*Buy Vip* and *Vente Privee*) which were pointed out by Quale because of the lack of consensus agreement in collected answers. Notice that this important issue is not taken into account by F-TOPSIS. So, we can conclude that Quale helps us to make a finer and deeper analysis than F-TOPSIS.

6.3.2 *e-Service Quality Modeling*

In the light of the analysis made in the previous section, we proposed characterizing e-Service Quality by the model depicted in Fig 6.6. Two latent sub-dimensions of e-Service Quality are observed: (1) Utilitarian Quality and (2) Hedonic Quality.

In addition, there are three latent sub-dimensions (Website Quality; Offered Service; Security) of Utilitarian Quality. They are somehow correlated as it was shown in Fig 6.2. More deeply, it is worthy to note that Website Quality is usually described in terms of website design and contents. In addition, Offered Service depends on guarantee, offer, and customization of service. Security involves payment management, privacy and trust.

We would like to remark once again the fact that evaluations given by users of B2C websites are inherently imprecise and uncertain, as they are based on human perceptions which are inherently subjective. Therefore, the design and implementation of the model introduced above must be made carefully in order to become operative, dynamic and adaptive in nature. Thus, we have implemented the proposed model in the form of a hierarchical fuzzy system with two layers.

Firstly, we addressed the knowledge extraction and representation task from experts. We asked a panel of on-line marketing experts to characterize inputs and outputs as well as relating them through fuzzy IF-THEN rules. Then, these expert KBs were enhanced by adding induced knowledge. We applied data mining tools provided by GUAJE software in order to extract valuable knowledge from data coming out of the second on-line survey on B2C websites which was described in the previous section. The combination of expert and induced knowledge was made by following HILK fuzzy modeling methodology (briefly introduced in Sect. 6.2.4 and implemented by GUAJE). It is worthy to note that the inference process is performed with the usual min-max fuzzy inference mechanism. Moreover, the well-known center of gravity is applied in the defuzzification stage.

We started with setting up a preliminary expert KB to assess e-Service Quality at the top of the hierarchy. It takes two input variables (Utilitarian Quality and Hedonic Quality) and produces one output variable (e-Service Quality). All the three variables are defined by strong fuzzy partitions (see pictures on the left of

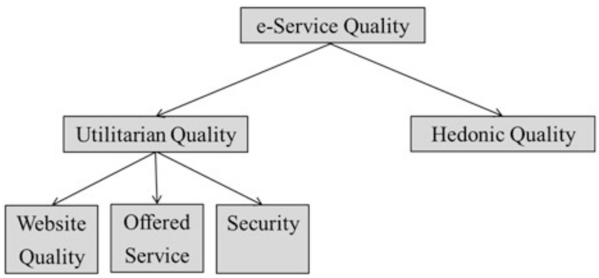

Fig. 6.6 Model for characterizing e-service quality

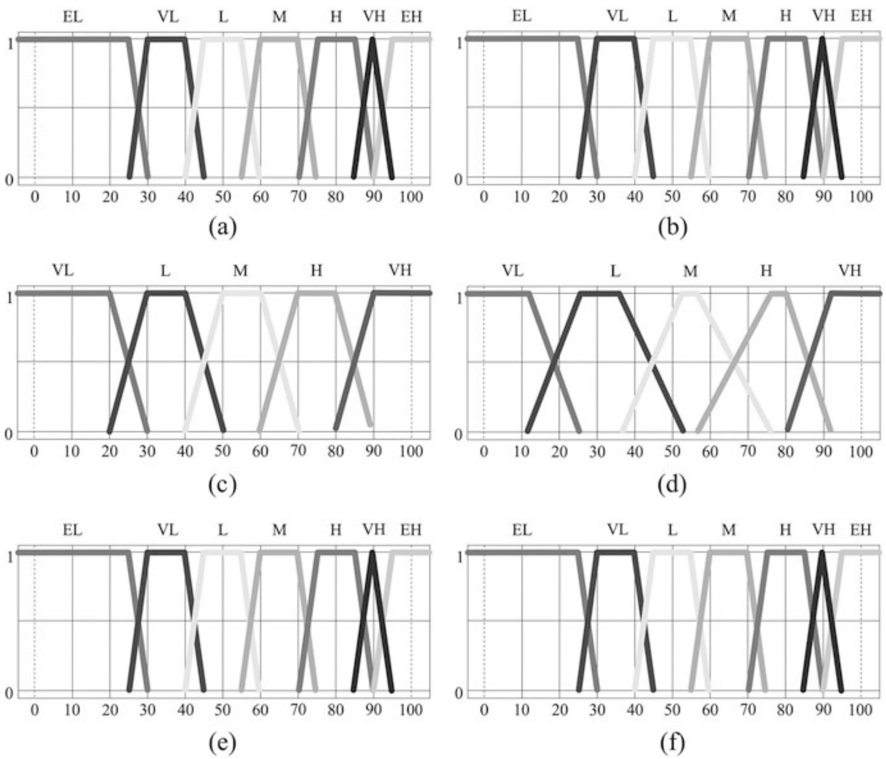

Fig. 6.7 Strong fuzzy partitions. (**a**) Expert utilitarian quality. (**b**) Optimized utilitarian quality. (**c**) Expert hedonic quality. (**d**) Optimized hedonic quality. (**e**) Expert e-service quality. (**f**) Optimized e-service quality

Fig. 6.7). e-Service Quality and Utilitarian Quality are made up of seven fuzzy sets each, with their related linguistic terms: Extremely Low (EL); Very Low (VL); Low (L); Medium (M); High (H); Very High (VH); Extremely High (EH). Hedonic Quality includes only five fuzzy sets (and the set of linguistic terms is a subset of the previous one). As it can be appreciated in Table 6.2, there are 35 expert rules. For example, the first rule can be read as follows "If Utilitarian Quality is Extremely Low and Hedonic Quality is Very Low Then e-Service Quality is Extremely Low".

Then, we applied the data mining tools provided by GUAJE in order to enrich the previous expert KB with knowledge automatically extracted from data. The adjustment of learning parameters and goodness of the designed KB was evaluated through 10-fold cross-validation. For each fold, we first derived rules from a pruned fuzzy decision tree. Secondly, we merged expert and induced rules through a linguistic simplification procedure. Finally, we refined fuzzy partitions by means of the Solis-Wetts tuning mechanism.

On the one hand, regarding training data, coverage measure achieved 100%, mean absolute error (MAE) was 1.47, and root mean square error (RMSE) was

Table 6.2 Expert rules to assess e-service quality in the textile and fashion sector

		Hedonic quality				
		Very Low (VL)	Low (L)	Medium (M)	High (H)	Very High (VH)
Utilitarian quality	Extremely Low (EL)	EL	VL	VL	L	M
	Very Low (VL)	VL	VL	L	M	M
	Low (L)	VL	L	L	M	H
	Medium (M)	VL	L	M	H	H
	High (H)	L	M	M	H	VH
	Very High (VH)	L	M	H	H	EH
	Extremely High (EH)	M	M	H	VH	EH

1.92. On the other hand, coverage arose to 100% while MEA was 1.68 and RMSE was 2.19, with respect to test data. Regarding interpretability indicators, in average, the number of rules was 11.5, the total rule length was 21, and the number of simultaneously fired rules was 2.88 in training and 2.78 in test.

Later, we repeated the same procedure to build the KB in the second layer of the hierarchy. It takes three input variables (Website Quality, Offered Service, and Security) and produces one output variable (Utilitarian Quality). Its goodness was also evaluated through 10-fold cross-validation. To sum up with, coverage was 99.91%, MAE was 2.43, and RMSE was 3.06, with respect to training data; while coverage was 100%, MAE was 2.36, and RMSE was 3.08, with respect to test data. In addition, the number of rules was 7, the total rule length was 12.6, and the number of simultaneously fired rules was 2.9 in training and 2.86 in test.

As a result, the designed model exhibits a good interpretability-accuracy trade-off since it is able to achieve high accuracy with a small set of highly readable linguistic rules. The final model considers all available data in combination with expert knowledge. Pictures on the right hand side of Fig. 6.7 depict the optimized fuzzy partitions. Moreover, the final 11 rules related to e-Service Quality assessment are as follows:

IF *Utilitarian Quality is EL OR VL* AND *Hedonic Quality is L OR M* THEN *e-Service Quality is VL*
IF *Utilitarian Quality is EL OR VL* AND *Hedonic Quality is M OR H* THEN *e-Service Quality is L*
IF *Utilitarian Quality is L* AND *Hedonic Quality is M* THEN *e-Service Quality is L*
IF *Utilitarian Quality is L OR M* AND *Hedonic Quality is H* THEN *e-Service Quality is M*
IF *Utilitarian Quality is M OR H* AND *Hedonic Quality is M* THEN *e-Service Quality is M*
IF *Utilitarian Quality is L OR M OR H* AND *Hedonic Quality is VH* THEN *e-Service Quality is H*
IF *Utilitarian Quality is M OR H OR VH* AND *Hedonic Quality is H* THEN *e-Service Quality is H*
IF *Utilitarian Quality is H OR VH* AND *Hedonic Quality is VH* THEN *e-Service Quality is VH*
IF *Utilitarian Quality is VH OR EH* AND *Hedonic Quality is VH* THEN *e-Service Quality is EH*
IF *Hedonic Quality is VL* THEN *e-Service Quality is VL*
IF *Hedonic Quality is L* THEN *e-Service Quality is L*

Once the proposed fuzzy model was validated, we embedded it in the core of an intelligent virtual assessor able to replicate the evaluations collected through the second survey described in the previous section. In practice, given the numerical values related to all factors defining a website (design, contents, guarantee, and so on), the virtual assessor is able to carry out a fuzzy inference yielding as result a global e-Service Quality score.

This way, the related ranking (with computed scores in brackets) is as follows: (1) Zara [82.5], (2) Asos [75.71], (3) eBay [72.47], (4) Privalia [70.46], (5) El Corte Inglés [69.22], and (6) Buy Vip [62.97]. It is worthy to note that Vente Privee was deliberately excluded from this ranking because, as we explained in the previous section, there was a lack of consensus among collected answers for the related website. Even though there are some minor differences between inferred scores and actual ones, this final ranking is fully in accordance with the one provided in the previous section (see Fig. 6.4). In consequence, the virtual assessor is ready to be used in prospective market research studies with the aim of estimating the e-Service Quality related to other websites different from those considered here; without requiring to disturb consumers with additional surveys.

6.4 Concluding Remarks and Future Work

This paper has presented a novel and efficient methodology for predictive analytics supported by business intelligence tools. We have expanded and further explored the knowledge on e-service quality, addressing a joint application of evaluations on hedonic and utilitarian dimensions by means of combined use of marketing methods (questionnaires) and the Computational Theory of Perceptions (Fuzzy Logic). Moreover, we have applied the paradigm of interpretable fuzzy modeling to deal properly with the uncertainty and imprecision characteristics of human perceptions.

As a result, we have translated sensory data collected through fuzzy rating scale-based questionnaires into valuable knowledge for business decision-making support. Moreover, the interpretability of the designed models is in the core of our human-centric approach. Accordingly, it yields reports easy to understand even by non-experts in the domain of interest as we have proved in a case study regarding B2C websites in the Spanish textile and fashion sector. Reports include several graphs easy to interpret along with a global ranking of retailers regarding e-service quality. Notice that the novel method presented in this paper is able to carry out a finer and deeper analysis than the well-known F-TOPSIS ranking method which we considered for comparison purposes. It is also worthy to remark that the designed virtual assessor is ready to automatically evaluate (without needing to ask directly to consumers) unknown websites out of the seven retailers under study.

In this work we have shown some of the main advantages and drawbacks of our fuzzy approach for e-service quality modeling. Fuzzy sets and systems are well-known because of their ability to properly handle imprecision and uncertainty.

Moreover, we adopted a human-centric modeling approach which yields a good interpretability-accuracy trade-off.

Nevertheless, a lot of work still remains to do. This paper opens the door to very challenging future research. For instance, the use of virtual assessors for reducing costs (mainly time and money) in future market research studies. Also, we plan exploring how to enhance our framework with advanced cloud computing and social network analysis tools.

Finally, let us remark that this work has been developed with the help of several software tools. Please, the interested reader is kindly referred to [1] for further details about them as well as other interesting fuzzy systems software.

Acknowledgements This work is supported by RYC-2016-19802 (Ramón y Cajal contract), and two MINECO projects TIN2017-84796-C2-1-R (BIGBISC) and TIN2014-56633-C3-3-R (ABS4SOW). All of them funded by the Spanish "Ministerio de Economía y Competitividad". Financial support from the Xunta de Galicia (Centro singular de investigación de Galicia accreditation 2016–2019) and the European Union (European Regional Development Fund - ERDF), is gratefully acknowledged.

References

1. Alcala-Fdez, J., & Alonso, J. M. (2016). A survey of fuzzy systems software: Taxonomy, current research trends and prospects. *IEEE Transactions on Fuzzy Systems, 24*(1), 40–56.
2. Alonso, J. M., Castiello, C., & Mencar, C. (2015). Interpretability of fuzzy systems: Current research trends and prospects. In *Handbook of computational intelligence* (pp. 219–237). Berlin: Springer.
3. Alonso, J. M, Castiello, C., & Mencar, C. (2018). A bibliometric analysis of the explainable artificial intelligence research field. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems - Theory and Foundations*. Communications in computer and information science (Vol. 853, pp. 3–15). Cham: Springer. https://doi.org/10.1007/978-3-319-91473-2_1
4. Alonso, J. M., & Magdalena, L. (2011). HILK++: An interpretability-guided fuzzy modeling methodology for learning readable and comprehensible fuzzy rule-based classifiers. *Soft Computing, 15*, 1959–1980.
5. Alonso, J. M., Pancho, D. P., Magdalena, L., Nunez, D. A., Sanchez, D. S., Suarez, P. F., et al. (2015). QUALE®: A new toolbox for quantitative and qualitative analysis of human perceptions. In *Proceedings of 16th World Congress of the International Fuzzy Systems Association (IFSA) 9th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT)* (pp. 659–666). Paris: Atlantis Press.
6. Carifio, J., & Perla, R. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences, 3*(3), 106–116.
7. Casillas, J., Cordon, O., Herrera, F., & Magdalena, L. (2003). *Interpretability issues in fuzzy modeling. Studies in fuzziness and soft computing*. Berlin: Springer.
8. Castro-Lopez, A., Puente, J., & Vazquez-Casielles, R. (2017). Fuzzy inference suitability to determine the utilitarian quality of B2C websites. *Applied Soft Computing, 57*, 132–143.
9. Chen, C. T., Lin, C. T., & Huang, S. F. (2006). A fuzzy approach for supplier evaluation and selection in supply chain management. *International Journal of Production Economics, 102*, 289–301.

10. CNMC. (2017). *Second quarter 2017 report on Spanish e-commerce*. Accessed May 2018.
11. de la Rosa de Saa, S., Gil, M. A., Gonzalez-Rodriguez, G., Lopez, M. T., & Lubiano, M. A. (2015). Fuzzy rating scale-based questionnaires and their statistical analysis. *IEEE Transactions on Fuzzy Systems, 23*, 111–126.
12. Gil, M. A., Lubiano, M. A., de la Rosa de Saa, S., & Sinova, B. (2015). Analyzing data from a fuzzy rating scale-based questionnaire. A case study. *Psicothema, 27*(2), 182–191.
13. Kabir, G., Sadiq, R., & Tesfamariam, S. (2014). A review of multi-criteria decision-making methods for infrastructure management. *Structure and Infrastructure Engineering, 10*, 1176–1210.
14. Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 1–55.
15. Pancho, D. P., Alonso, J. M., & Magdalena, L. (2013). Quest for interpretability-accuracy trade-off supported by fingrams into the fuzzy modeling tool GUAJE. *International Journal of Computational Intelligence Systems, 6*(Suppl. 1), 46–60.
16. Parasuraman, A., Zeithaml, V., & Malhotra, A. (2005). E-S-Qual: A multiple-item scale for assessing electronic service quality. *Journal of Service Research, 7*(10), 1–21.
17. Posner, H. (2016). *Fashion marketing*. Barcelona: Gustavo Gili Editorial.
18. Quirós, P., Alonso, J. M., & Pancho, D. P. (2016). Descriptive and comparative analysis of human perceptions expressed through fuzzy rating scale-based questionnaires. *International Journal of Computational Intelligence Systems, 9*(3), 450–467.
19. Rares, O. (2014). Measuring perceived service quality offline vs. online: A new PeSQ conceptual model. *Procedia Economics and Finance, 15*, 538–551.
20. Salehi, M., Salimi, M., & Haque, A. (2013). The impact of online customer experience (OCE) on service quality in Malaysia. *Word Applied Sciences Journal, 21*(11), 1621–1631.
21. Sun, C. C., & Lin, G. T. R. (2009). Using fuzzy TOPSIS method for evaluating the competitive advantages of shopping websites. *Expert Systems with Applications, 36*(9), 11764–11771.
22. The European Apparel and Textile Confederation. (2016). *Taking action for the future of the European textile and clothing industry*. Accessed May 2018.
23. Wang, Y. J., & Lee, H. (2007). Generalizing TOPSIS for fuzzy multiple-criteria group decision-making. *Computers and Mathematics with Applications, 53*, 1762–1772.
24. Zadeh, L. A. (2001). A new direction in AI: Toward a computational theory of perceptions. *Artificial Intelligence Magazine, 22*(1), 73–84.
25. Zeng, X., Ruan, D., & Koehl, L. (2008). Intelligent sensory evaluation: Concepts, implementations and applications. *Mathematics and Computers in Simulation, 77*, 443–452.